AMENDMENTS TO THE SPECIFICATION

Amend the paragraph found in the Summary of the Invention on page 3, lines 1-3 as follows:

In accordance with the various exemplary aspects of this invention, log-linear models are used wherein parameters may be trained with sparse or incomplete training data. The log-linear function receives the multitude of speech features obtained from the input speech and determines a posterior probability of each of a plurality of hypothesized linguistic units given the extracted multitude of speech features by applying the formula:

$$P(H_j \mid features) = P(w_1^k \mid o_1^T) = \prod_{i=1}^{k} P(w_i \mid w_1^{i-1}, o_1^T),$$

where:

H_j is a jth hypothesis that contains a sequence of word (or other linguist unit) sequence w_1^k = w1w2…wk

i is an *index pointing to the ith word (or unit)*

k is a *number of words (units) in the hypothesis*

T is a *length of the speech signal (e.g. number of frames)*

w_1^k is a sequence of words associated with the hypothesis H_j, and

o_1^T is a sequence of acoustic observations.

with the conditional probabilities represented by a maximum entropy log-linear model:

$$P(w_i \mid w_1^{i-1}, o_1^T) = \frac{e^{\sum_j \lambda_j f_j(w_i, w_1^{i-1}, o_1^T)}}{Z(w_1^{i-1}, o_1^T)},$$

where:

λ_j are parameters of the log-linear model,

f_j are a multitude of features extracted,

and

Z is a normalization factor that ensures that Equation 2 is a true probability (will sum up to 1).